

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2004-287477

(43)Date of publication of application : 14.10.2004

(51)Int.Cl. G06F 3/06
G06F 12/00

(21)Application number : 2003-075187

(71)Applicant : HITACHI LTD

(22)Date of filing : 19.03.2003

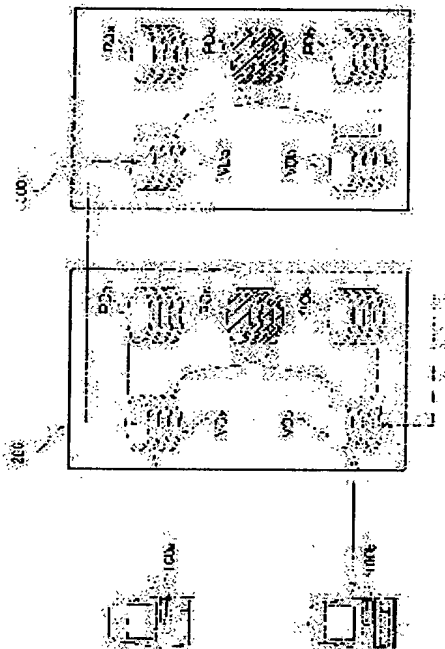
(72)Inventor : KIMURA SHINJI
OSHIMA SATOSHI
HASHIMOTO TAKASHI

(54) DATA STORAGE/READOUT CONTROL FOR MASS STORAGE DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To improve data readout speed from a disk device in a computer system capable of transferring data through a node device 200 between a client 100 and a storage device 300.

SOLUTION: A common volume PDc and unique volumes PDa, PDb are defined on a disk device of a storage device 300. Data common to each client are stored in the common volume, and data unique for every client are stored in the unique volume. When each client requests readout of data to virtual volumes VDa, VDb, the storage device reads corresponding data out of the common volume or unique volume. Consequently, even if access concentration from a plurality of clients is occurred, a lowering of readout speed can be suppressed since data are mainly read out of the common volume.



LEGAL STATUS

[Date of request for examination] 04.10.2005

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-287477

(P2004-287477A)

(43) 公開日 平成16年10月14日(2004. 10. 14)

(51) Int. Cl.⁷

G06F 3/06

G06F 12/00

F 1

G06F 3/06

301Z

テーマコード (参考)

5B065

G06F 3/06

304H

5B082

G06F 12/00

514E

G06F 12/00

545A

審査請求 未請求 請求項の数 13 O L (全 19 頁)

(21) 出願番号 特願2003-75187 (P2003-75187)
 (22) 出願日 平成15年3月19日(2003. 3. 19)

(71) 出願人 000005108
 株式会社日立製作所
 東京都千代田区神田駿河台四丁目6番地
 (74) 代理人 110000028
 特許業務法人明成国際特許事務所
 (72) 発明者 木村 信二
 神奈川県川崎市麻生区王禅寺1099番地
 株式会社日立製作所システム開発研究所
 内
 (72) 発明者 大島 訓
 神奈川県川崎市麻生区王禅寺1099番地
 株式会社日立製作所システム開発研究所
 内

最終頁に続く

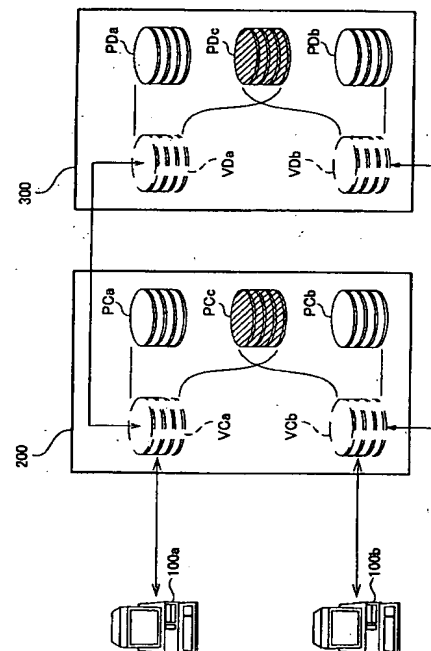
(54) 【発明の名称】 大容量記憶装置におけるデータの格納・読み出し制御

(57) 【要約】

【課題】 クライアント100およびストレージ装置300間で、ノード装置200を介してデータを授受可能な計算機システムにおいて、ディスク装置からのデータ読みだし速度を向上する。

【解決手段】 ストレージ装置300のディスク装置に、共通ボリュームPDcと固有ボリュームPDa、PDbを定義する。共通ボリュームには各クライアントに共通のデータ、固有ボリュームにはクライアントごとに固有のデータを格納する。各クライアントが、仮想ボリュームVDa、VDbに対してデータの読み出しを要求すると、ストレージ装置は、共通ボリュームまたは固有ボリュームから対応するデータを読み出す。こうすることにより、複数クライアントからのアクセスが集中した場合でも、共通ボリュームからのデータの読み出しが主となるため、読みだし速度の低下を抑制することができる。

【選択図】 図2



【特許請求の範囲】

【請求項 1】

ディスク装置に記憶されたデータを複数のクライアントからの要求に応じて提供するデータ提供装置であって、

前記ディスク装置は、

前記複数のクライアントが共通して使用する所定容量の共通ボリュームと、

前記複数のクライアントに対応付けて区切られた複数の固有ボリュームとを有し、

前記複数のクライアントごとに個別に仮想的なボリュームを割り当てる仮想ボリューム管理部と、

該仮想的なボリューム内に仮想的に定義された仮想ブロックと、前記共通ボリュームおよび固有ボリュームについて物理的に規定された各物理ブロックとの対応づけを管理するブロック対応管理部と、

前記クライアントから前記仮想ブロックの指定を伴う読出コマンドを受け付けるコマンド受付部と、

前記ブロック対応管理部を参照して、該仮想ブロックに対応するデータを、該クライアントに対応する固有ボリュームまたは前記共通ボリュームから読み出し、該読出コマンドを発信したクライアントに対して出力する読出制御部とを備えるデータ提供装置。

【請求項 2】

請求項 1 記載のデータ提供装置であって、

前記クライアントは、該データ提供装置からのデータ供給を受けてリモートブート可能であり、

前記共通ボリュームは、該リモートブート時に前記複数のクライアントに供給すべきデータを格納したデータ提供装置。

【請求項 3】

請求項 1 記載のデータ提供装置であって、

前記ブロック対応管理部は、前記複数の仮想ボリュームにおける少なくとも一部の仮想ブロックを、同一の共通ブロックに対応づけるデータ提供装置。

【請求項 4】

請求項 1 記載のデータ提供装置であって、

該データ提供装置と前記複数のクライアントとの間には、データのキャッシュ機能を備える複数のノード装置が接続されており、

該データ提供装置は、

前記ノード装置からの要求に応じて、各データが共通ボリュームおよび固有ボリュームのいずれに格納されているかを特定するための特定情報を該ノード装置に出力する特定情報出力部を備えるデータ提供装置。

【請求項 5】

請求項 4 記載のデータ提供装置であって、

前記特定情報は、前記仮想ブロックと前記物理ブロックとの対応関係であるデータ提供装置。

【請求項 6】

請求項 4 記載のデータ提供装置であって、

前記特定情報出力部は、前記読出コマンドに関わらず所定のタイミングで、複数の仮想ブロックに関する前記特定情報を出力するデータ提供装置。

【請求項 7】

請求項 4 記載のデータ提供装置であって、

前記ノード装置からの要求は、仮想ブロックの指定を含んでおり、

前記特定情報出力部は、該指定された仮想ブロックに関する前記特定情報を出力するデータ提供装置。

【請求項 8】

請求項 1 記載のデータ提供装置であって、

10

20

30

40

50

所定の上位データ提供装置と前記複数のクライアントに接続され、両者間でのデータの授受を中継するノード装置として構成されており、

前記クライアントからの読出コマンドで指定された仮想ブロックに対応するデータが、前記ディスク装置に格納されていない場合には、該データの読出コマンドを前記上位データ提供装置に送信するコマンド送信部と、

前記上位データ提供装置から、該データ、および該データが共通ボリュームおよび固有ボリュームのいずれに格納されていたデータであるかを表す特定情報を受信するデータ受信部と、

該受信したデータを前記クライアントに転送する転送部と、

該受信したデータを前記特定情報に応じて前記共通ボリュームおよび固有ボリュームのいずれかに格納するとともに、該格納場所に応じて前記ブロック対応管理部における前記仮想ブロックと前記物理ブロックとの対応づけを更新する格納制御部とを備えるデータ提供装置。

10

【請求項 9】

請求項 8 記載のデータ提供装置であって、

前記特定情報は、前記仮想ブロックと他の仮想ブロックとのデータの共通性も特定可能であり、

前記コマンド送信部による読出コマンドの送信に先立って、該特定情報の出力を前記上位データ提供装置に要求する特定情報要求部と、

前記コマンド送信部は、該出力要求に応じて該上位データ提供装置から受信した前記特定情報に基づいて、前記データが、前記ディスク装置に格納されているか否かを判断するデータ提供装置。

20

【請求項 10】

請求項 8 記載のデータ提供装置であって、

前記特定情報は、前記仮想ブロックと他の仮想ブロックとのデータの共通性も特定可能であり、

該特定情報を前記上位データ提供装置から取得し、管理する特定情報管理部を備え、

前記コマンド送信部は、該特定情報管理部を参照して、前記データが、前記ディスク装置に格納されているか否かを判断するデータ提供装置。

【請求項 11】

30

請求項 10 記載のデータ提供装置であって、

前記特定情報管理部は、前記特定情報として、前記仮想ブロックと前記上位データ提供装置における物理ブロックとの対応関係を管理しており、

前記ブロック対応管理部は、

前記上位データ提供装置における物理ブロックと、該データ提供装置における物理ブロックとの対応関係を管理する物理ブロック管理部を有し、

前記特定情報管理部および物理ブロック管理部を参照して、前記仮想ブロックと該データ提供装置における物理ブロックとの対応関係を特定するデータ提供装置。

【請求項 12】

ディスク装置に記憶されたデータを複数のクライアントからの要求に応じて提供するデータ提供装置の制御方法であって、

40

前記ディスク装置について、前記複数のクライアントが共通して使用する所定容量の共通ボリュームと、前記複数のクライアントに対応付けて区切られた複数の固有ボリュームとを定義する設定情報を管理するステップと、

前記複数のクライアントごとに個別に仮想的なボリュームを割り当てるステップと、

該仮想的なボリューム内に仮想的に定義された仮想ブロックと、前記共通ボリュームおよび固有ボリュームについて物理的に規定された各物理ブロックとの対応づけを管理するステップと、

前記クライアントから前記仮想ブロックの指定を伴う読出コマンドを受け付けるステップと、

50

前記ブロック対応管理部を参照して、該仮想ブロックに対応するデータを、該クライアントに対応する固有ボリュームまたは前記共通ボリュームから読み出し、該読出コマンドを発信したクライアントに対して出力するステップとを備えるデータ提供装置の制御方法。

【請求項13】

ディスク装置に記憶されたデータを複数のクライアントからの要求に応じて提供するデータ提供装置の動作を制御するためのコンピュータプログラムであって、

前記ディスク装置について、前記複数のクライアントが共通して使用する所定容量の共通ボリュームと、前記複数のクライアントに対応付けて区切られた複数の固有ボリュームとを定義する設定情報を管理する機能と、

前記複数のクライアントごとに個別に仮想的なボリュームを割り当てる機能と、

該仮想的なボリューム内に仮想的に定義された仮想ブロックと、前記共通ボリュームおよび固有ボリュームについて物理的に規定された各物理ブロックとの対応づけを管理する機能と、

前記クライアントから前記仮想ブロックの指定を伴う読出コマンドを受け付ける機能と、前記ブロック対応管理部を参照して、該仮想ブロックに対応するデータを、該クライアントに対応する固有ボリュームまたは前記共通ボリュームから読み出し、該読出コマンドを発信したクライアントに対して出力する機能とを該データ提供装置において実現するためのコンピュータプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ストレージ装置、ノード装置など、複数のクライアントからアクセスを受ける大容量記憶装置におけるデータの格納・読み出し制御に関する。

【0002】

【従来の技術】

企業などでは、複数のクライアントコンピュータ（以下、単にクライアントと称する）をネットワークで接続した大規模な計算機システムが構築されている。このような計算機システムでは、各クライアントで業務を遂行する際に使用する種々のデータやプログラムを提供するため、ストレージ装置と呼ばれる大容量の記憶装置が接続されている。ストレージ装置は、大容量のハードディスクを備えており、各クライアントのユーザは、自己に割り当てられた領域、または共有の領域に対して、データやプログラムなどの格納・読み出しを行う。通常、ストレージ装置へのアクセスには、SCSI (Small Computer System Interface) と呼ばれるプロトコルが使用される。近年では、SCSI用のコマンドを、いわゆるIPパケットに分割してネットワーク上を転送するためのiSCSIと呼ばれるプロトコルが提案されている。iSCSIを用いることにより、従来よりも更に多数のクライアント、および遠隔地のクライアントがストレージ装置にアクセス可能な環境が提供されつつある。

【0003】

ストレージ装置の活用方法の一つとして、リモートブートと呼ばれるクライアントの起動方法が挙げられる。リモートブートでは、起動に必要なオペレーティングシステム、環境設定ファイルなどの各種ファイルが、予めストレージ装置の各ユーザ用の領域に格納されている。各クライアントは、ネットワークを介してこれらのファイルをストレージ装置から読み込むことにより、起動することができる。

【0004】

リモートブートする場合、クライアントは、ストレージ装置へのアクセスを実現するための比較的小さなブートプログラムをROMに記憶していれば十分であり、ハードディスクを備えている必要がない。ハードディスクに格納されるべきデータはストレージ装置で一元管理されるから、仮にクライアントコンピュータが破損、盗難などに遭っても、重要なデータの遺失、漏洩を逃れることができるというセキュリティ上の利点がある。また、各ユーザは、ネットワーク上に接続されたいずれのクライアントも、一定の環境設定で起動

させることができるという実用上の利点もある。リモートブートについては、例えば、特許文献1～3に開示されている。

【0005】

【特許文献1】

特開平6-332716号公報

【特許文献2】

特開2001-75853号公報

【特許文献3】

特開2000-259583号公報

【0006】

【発明が解決しようとする課題】

リモートブートを行う場合、多数のクライアントがほぼ同時に起動すると、ストレージ装置へのアクセスが集中し、各クライアントへのファイル供給が遅れ、ひいては起動に長時間かかるという状況を招くことがあった。本願発明者の調査によれば、10台程度のクライアントが同時に起動した場合には、既にユーザが支障を感じる程度に起動時間が長引くことが見いだされている。クライアントの起動は、通常、企業の始業時に集中する傾向が強いので、このような状況は比較的起きやすかった。

【0007】

ストレージ装置へのアクセスを緩和するため、クライアントとストレージ装置との間に、キャッシュ用のメモリを備えたノード装置を設けることがある。かかるシステムでは、ノード装置にキャッシュ済みのデータは、ストレージ装置にアクセスするまでなくクライアントに供給可能であるため、ストレージ装置へのアクセスが緩和される。しかし、かかるシステムにおいても、複数のクライアントが同時に起動する場合には、ノード装置へのアクセスが集中するため、起動時間が長くなるという課題は、十分に解決できなかった。

【0008】

上述の課題は、リモートブートに限るものではなく、またストレージ装置、ノード装置に限るものでもない。いわゆる複数のクライアントからアクセスを受け得る記憶装置におけるデータの読み出し時に共通の課題であった。本発明は、かかる課題に鑑み、複数のクライアントから記憶装置へのアクセスが集中した場合でも、記憶装置からの円滑なデータ供給を実現することを目的とする。

【0009】

【課題を解決するための手段】

本発明では、データ提供装置に次の構成を適用することで、上記課題の少なくとも一部の解決を図る。データ提供装置とは、ディスク装置に記憶されたデータを複数のクライアントからの要求に応じて提供する装置であり、例えば、ストレージ装置やノード装置などが含まれる。

【0010】

本発明のデータ提供装置に備えられたディスク装置は、複数のクライアントが共通して使用する所定容量の共通ボリュームと、複数のクライアントに対応付けて区切られた複数の固有ボリュームとを有する。データ提供装置は、これらの共通ボリューム、固有ボリュームとは別に、各クライアントに個別に仮想的なボリュームを割り当てる。この仮想的なボリューム内にはクライアントとのデータ授受で使用する仮想的に定義されたデータ格納場所として仮想ブロックが用意されている。データ提供装置は、この仮想ブロックと、共通ボリュームおよび固有ボリュームについて物理的に規定された各物理ブロックとの対応づけを管理している。

【0011】

本発明において、仮想ブロックの論理的な管理単位としてのクライアントは、例えば、データ提供装置に接続される計算機単位、各計算機を利用するユーザ単位、各ユーザの環境設定単位などの設定が可能である。計算機単位とする場合、各計算機を使用するユーザに関わらず、計算機ごとに一定の利用環境を提供することができる。ユーザ単位とすれば、

10

20

30

40

50

どの計算機を利用するかに関わらず、ユーザ固有の利用環境を提供することができる。環境設定単位とすれば、各ユーザは複数の利用環境を使い分けることができる。以下、本明細書において、「クライアント」という用語は、仮想ブロックの論理的な管理単位を意味するものとして用い、ハードウェアとしてのクライアントコンピュータに対しては「計算機」という用語を用いるものとする。

【0012】

クライアントからデータ提供装置へは、仮想ブロックを指定してデータの読出コマンドが送信される。データ提供装置は、上述の対応付けに基づき、この読出コマンドで指定された仮想ブロックに対応するデータを、クライアントに対応する固有ボリュームまたは共通ボリュームから読み出す。クライアントの識別は、計算機を特定するアドレス情報や、ユーザやその利用環境を特定するためのID情報によって行うことができる。

10

【0013】

本発明のデータ提供装置によれば、共通ボリュームと固有ボリュームに区分けして、データを格納することができる。このように区分けしてデータを格納していても、仮想ボリュームを用いることにより、クライアントは、共通ボリュームおよび固有ボリュームという区分を考慮することなく、データにアクセスすることができる。

【0014】

一般に、ディスク装置では、データを読み出すためのヘッドの物理的な移動量がデータの読み出し速度に少なからず影響を与える。本発明のデータ提供装置は、上述のデータ格納方法を採用することにより、複数のクライアントから共通して読み出されるデータを共通ボリュームでまとめて管理することができる。共通ボリュームにまとめて格納されたデータについては、複数のクライアントからのアクセスが集中しても、ヘッドの移動量を抑制することができる。データの読み出し速度低下を抑制することができる。本発明において、共通ボリュームは一つに限る必要はなく、複数設けても構わない。但し、読み出し速度確保という観点からは、共通ボリュームは一つにまとめておくことが好ましい。

20

【0015】

共通ボリュームには、複数のクライアントからのアクセスが集中する可能性のあるデータを格納することが好ましい。例えば、データ提供装置を用いて、クライアントのリモートブートを行う場合には、リモートブート時に各クライアントに供給すべきデータ、例えば、オペレーティングシステムなどのファイルを共通ボリュームに格納することができる。また、企業内で固有の業務を遂行するための専用システムでは、その業務専用のアプリケーションプログラムを共通ボリュームに格納してもよい。

30

【0016】

共通ボリュームには、必ずしも複数のクライアントに共通して提供されるデータのみを格納する必要はなく、いずれかのクライアントに固有のデータが混在していても構わない。例えば、企業内において出納業務を行うクライアントと、生産管理を行うクライアントが混在している場合を考える。各クライアントは、起動時に、出納業務用のアプリケーションおよび生産管理用のアプリケーションをそれぞれ起動させる。このような場合、これらのアプリケーションプログラムは始業時にそれぞれのクライアントから同時に読み出される可能性が高くなるから、共通ボリュームに格納しておくことが好ましい。

40

【0017】

本発明において、異なるクライアントの仮想ブロックを同一の共通ブロックに対応づけてもよい。こうすることにより、各クライアントは同一の共通ブロックからデータを読み出すことが可能となる。従って、複数のクライアントに共通して提供されるデータが存在する場合には、共通ボリュームに一組のデータを格納しておけば足り、複数のデータを重複して格納しておく必要がなくなる。

【0018】

本発明のデータ提供装置は、ノード装置を備える計算機システムに適用することも可能である。ノード装置とは、データ提供装置と複数のクライアントとの間に接続され、両者間でデータの中継を行う装置である。中継されるデータは、ノード装置のディスク装置にキ

50

キャッシュされる。この場合、キャッシュされたデータの読み出し速度の低下を抑制するためには、ノード装置においても、データ提供装置の共通ボリュームに格納されているデータと固有ボリュームに格納されているデータとを区分して取扱可能とすることが好ましい。

【0019】

ノード装置における上記取扱を可能とするため、データ提供装置は、例えば、各データが共通ボリュームおよび固有ボリュームのいずれに格納されているかを特定するための特定情報をノード装置に出力可能としてもよい。特定情報は、例えば、仮想ブロックと物理ブロックとの対応関係とすることができる。共通ボリュームおよび固有ボリュームのいずれに格納されているかを表すフラグを、データごとに特定情報として添付するようにしてもよい。

10

【0020】

データ提供装置がノード装置に特定情報を送出するタイミングは、種々の設定が可能である。例えば、ノード装置の起動時など、読出コマンドに関わらない所定のタイミングとしてもよい。この場合、複数の仮想ブロックに関する特定情報をまとめて出力することが好ましい。こうすることにより、ノード装置は、クライアントからの読出コマンドの受信時に、データが共通ボリューム、固有ボリュームのいずれに格納されているかを確認する必要がなくなるため、読み出し時の処理を簡略化することができる。

【0021】

別の態様として、ノード装置からデータ提供装置に、仮想ブロックを指定して特定情報の送出を要求したタイミングとしてもよい。データ提供装置は、この要求に応じて、指定された仮想ブロックに関する特定情報を送出する。特定情報のみを送出してもよいし、データとともに送出手でもよい。かかる態様では、ノード装置は、キャッシュされていないデータについての特定情報を管理する必要がなくなるため、ハードウェア資源の節約、管理負荷の軽減を図ることができる。

20

【0022】

特定情報は、各クライアントの仮想ブロックと物理ブロックとを対応づける情報であるから、クライアントに依存した情報である。データ提供装置は、あるクライアントに対応した特定情報を、そのクライアントが利用するノード装置に対してのみ送出手でもよいし、クライアントが利用するか否かに関わらず複数のノード装置に送出手するようにしてもよい。前者によれば、無駄な通信を省きトラフィックの軽減を図ることができる。また、ノード装置にとっても無用な特定情報を管理する必要がなくなるという利点がある。後者によれば、ユーザが別のノード装置に接続されている計算機に移動した場合でも、従前の特定情報を活用することができる利点がある。

30

【0023】

本発明は、ノード装置として構成することもできる。ノード装置には、データ提供装置とクライアントの間でデータを中継するという機能を有する点を除けば、基本的にデータ提供装置と同等の構成を適用することができる。この意味で、本発明のノード装置は、データ提供装置の一形態、即ち、上位のデータ提供装置からデータの供給を受け、クライアントにデータを提供する機能を有するデータ提供装置としてとらえることもできる。以下、説明の便宜上、データの供給源となる装置をデータ提供装置、中継機能を備える装置をノード装置と称し、ノード装置における物理ブロックを「キャッシュブロック」と称する。

40

【0024】

本発明としてのノード装置は、データ提供装置と同様の構成を適用することにより、クライアントから指定された仮想ブロックに対応するデータを、共通ボリュームまたは固有ボリュームから読み出すことができる。また、ノード装置は、データ提供装置から受信したデータを、共通ボリュームと固有ボリュームに分けてディスク装置内にキャッシュする。

【0025】

このようなキャッシュ機能は次の構成によって実現される。ノード装置は、クライアントから指定されたデータが、ノード装置のディスク装置に格納されていない場合には、その

50

データの読出コマンドをデータ提供装置に送信する。データ提供装置からは、指定したデータ、及びそのデータが共通ボリュームおよび固有ボリュームのいずれに格納されていたデータであるかを表す特定情報を受信する。ノード装置は、受信したデータをクライアントに転送するとともに、特定情報に応じて共通ボリュームおよび固有ボリュームのいずれかに格納する。また、格納先に応じて、ノード装置における仮想ブロックとキャッシュブロックとの対応づけを更新する。

【0026】

特定情報は、格納先だけでなく、仮想ブロックと他の仮想ブロックとのデータの共通性も特定可能とすることが好ましい。こうすれば、ノード装置は、各クライアントからの読出コマンドで指定された仮想ブロック間の共通性を判断することができ、既にキャッシュされているか否かの判断に活用することができる。また、既にキャッシュ済みのデータについては、キャッシュを省略することが可能となり、ノード装置のハードウェア資源の節約を図ることも可能となる。

【0027】

このような特定情報の取得は、種々のタイミングで行うことができる。例えば、ノード装置からデータ提供装置への読出コマンドの送信に先立って、特定情報の出力をデータ提供装置に要求してもよい。こうすれば、データの受信前にキャッシュ済みか否かを確認することができ、トラフィックの軽減を図ることができる。こうして取得された特定情報をノード装置内で管理し、以後、キャッシュ済みか否かの判断に活用してもよい。特定情報をデータ提供装置から事前に取得して管理しておいてもよい。

【0028】

ここで、特定情報としては、仮想ブロックとデータ提供装置における物理ブロックとの対応関係を用いることができる。この対応関係は、データ提供装置がデータの格納場所を管理するために保有している情報をそのまま活用することができる。ノード装置では、この物理ブロックとキャッシュブロックとの対応関係を管理することにより、間接的に仮想ブロックとキャッシュブロックを対応づけることができる。こうすることにより、データの格納、キャッシュ済みか否かの判断に、データ提供装置が有する情報を活用することができる。

【0029】

本発明において、上述した種々の特徴は、適宜、組み合わせたり、一部を省略したりすることが可能である。また、本発明は、上述したデータ提供装置、ノード装置に限らず種々の態様で構成することが可能である。例えば、これらの装置を接続した計算機システムとして構成してもよいし、データ提供装置、ノード装置、計算機システムの制御方法として構成してもよい。コンピュータによりこの制御を実現するためのコンピュータプログラムおよびこれを記録した記憶媒体として構成してもよい。ここで、記憶媒体としては、フレキシブルディスクやCD-ROM、DVD、光ディスク、ICカード、ROMカートリッジ、パンチカード、バーコードなどの符号が印刷された印刷物、コンピュータの内部記憶装置（RAMやROMなどのメモリ）および外部記憶装置などコンピュータが読取り可能な種々の媒体を利用できる。

【0030】

【発明の実施の形態】

本発明の実施の形態について以下の順序で説明する。

- A. システム構成；
- B. データの格納状態；
- C. 機能ブロック；
- D. テーブルのデータ構造；
- E. データ読出処理；
- F. 変形例；

【0031】

- A. システム構成；

図1は実施例としての計算機システムの構成を示す説明図である。この計算機システムでは、ストレージ装置300に、複数のノード装置200および計算機100がIPパケットを授受可能なネットワークで接続されている。ネットワークとしては、LAN (Local Area Network)、イントラネット、インターネットなどを適用することができる。

【0032】

ストレージ装置300は、図示するハードウェア構成を有するコンピュータとして構成されており、大容量のディスク装置310に蓄積されたデータをノード装置200経由で各計算機100に提供することができる。また、各計算機100はハードディスク310にデータを格納することもできる。NIC (Network Interface Card) 304は、ネットワークを介してデータを授受するためのインタフェースである。I/O 305は、入出力装置用のインタフェースである。

【0033】

ストレージ装置300では、CPU 301が、ディスク装置310およびROM 303に記憶された制御用のプログラムに従って、データの授受を制御する。メモリ302は、このプログラムの動作時に使用される主記憶である。

【0034】

ノード装置200は、図示するハードウェア構成を有するコンピュータとして構成されており、ストレージ装置300と計算機100の間で授受されるデータを中継する機能を奏する。ディスク装置210およびメモリ202は、中継されるデータを一時的に蓄積するキャッシュのために利用される。NIC (Network Interface Card) 204は、ネットワークを介してデータを授受するためのインタフェースである。I/O 205は、入出力装置用のインタフェースである。

【0035】

ノード装置200では、CPU 201が、ディスク装置210およびROM 203に記憶された制御用のプログラムに従って、データの中継およびキャッシュを制御する。メモリ202は、このプログラムの動作時に使用される主記憶である。

【0036】

計算機100は、ハードディスクを備えないディスクレスコンピュータである。内部には、CPU 101、RAM 102、ROM 103、NIC 104およびI/O 105が備えられている。計算機100は、起動時には、必要なオペレーティングシステムなどのファイルをストレージ装置300から読み込むことで、リモートブートする。ROM 103には、リモートブート用のプログラムが予め記憶されている。

【0037】

ストレージ装置300、ノード装置200、計算機100の動作を制御するためのプログラムは、CD-ROMなどの記録媒体によってストレージ装置に供給してもよい。

【0038】

B. データの格納状態；

図2はデータの格納および読出状態を模式的に示す説明図である。ここでは、2台のクライアント100a、100bが接続されている状態を示した。クライアント100a、100bは、単にハードウェアとしての計算機というだけでなく、その計算機を使用するユーザが確定している状態を意味する。

【0039】

ストレージ装置300では、クライアント100a、100bのそれぞれに対応する仮想ボリュームVDa、VDbが定義されている。仮想ボリュームVDa、VDbとは、ストレージ装置300のディスク装置310内に設けられた物理的な領域ではなく、各クライアントのデータを管理するための論理的なディスクを意味する。ディスク装置310の物理的な領域には、クライアント100a、100bのそれぞれに固有に割り当てられる固有ボリュームPDa、PDbと、両者が共通して使用する共通ボリュームPDcがパーティションによって定義されている。図示する通り、仮想ボリュームVDaは固有ボリュー

10

20

30

40

50

ムPD a および共通ボリュームPD cに割り当てられており、仮想ボリュームVD bは固有ボリュームPD bおよび共通ボリュームPD cに割り当てられている。以下、仮想ボリュームVD a、VD bにおけるデータの格納場所を仮想ブロック、ディスク装置310内におけるデータの格納場所を物理ブロックと称する。

【0040】

ノード装置200においても同様に、クライアント100 a、100 bのそれぞれに対応する仮想ボリュームVC a、VC bが定義されている。ノード装置200のディスク装置210の物理的な領域には、クライアント100 a、100 bのそれぞれに固有に割り当てられる固有ボリュームPC a、PC bと、両者が共通して使用する共通ボリュームPC cが定義されている。図示する通り、仮想ボリュームVC aは固有ボリュームPC aおよび共通ボリュームPC cに割り当てられており、仮想ボリュームVC bは固有ボリュームPC bおよび共通ボリュームPC cに割り当てられている。以下、ノード装置200のディスク装置210におけるデータの格納場所をキャッシュブロックと称する。

10

【0041】

クライアント100 aは、リモートブート時など種々のタイミングで、仮想ブロックを指定してデータの読み出し要求をノード装置200に送信する。ノード装置200は、指定されたデータが仮想ボリュームVC aにキャッシュ済みであるか否かを確認する。キャッシュ済みである場合、ノード装置200は、仮想ボリュームVC aに対応する固有ボリュームPC aまたは共通ボリュームPC cからデータを読みだし、クライアント100 aに返信する。

20

【0042】

仮想ボリュームVC aにキャッシュされていない場合には、ノード装置200は、読出コマンドをストレージ装置300に送信する。ストレージ装置300は、読出コマンドで指定された仮想ブロックに対応するデータを、固有ボリュームPD aおよび共通ボリュームPD cから読み出し、ノード装置200に返信する。ノード装置200は、このデータをクライアント100 aに返信するとともに、仮想ボリュームVC aに対応する固有ボリュームPC aまたは共通ボリュームPC cに格納する。クライアント100 bからの読出コマンドについても同様の処理が行われる。

【0043】

図3はデータの格納例を示す説明図である。図の左側には、ストレージ装置300で管理される仮想ボリュームVD a、VD bに格納されているファイルを示し、右側には固有ボリュームPD a、PD bおよび共通ボリュームPD cに格納されているファイルを示した。仮想ボリュームVD aには、クライアント100 aで使用されるオペレーティングシステムOS a用のファイル、および種々のアプリケーションプログラムや文書データが格納されている。オペレーティングシステムOS aには図示する6つのファイルが含まれており、IPL、DOS、SYS、INOUT、SYS、COM、COMはユーザの環境設定等に依存しないファイル、CONF、SYS、AUTO、BATはデフォルトの環境設定を表すファイルであるものとする。仮想ボリュームVD bには、クライアント100 bで使用されるオペレーティングシステムOS b用のファイル等が格納されている。オペレーティングシステムOS bは、環境設定を表すファイルCONF、SYS b、AUTO、BAT bがオペレーティングシステムOS aと相違している。

30

40

【0044】

本実施例では、各クライアントに固有のファイルを固有ボリュームPD a、PD bに格納し、各クライアントが共通して使用するファイルを共通ボリュームPD cに格納する。従って、クライアント100 aが仮想ボリュームVD aに格納されているファイルのうち、アプリケーション1、アプリケーション2、および文書ファイルaは、実際には固有ボリュームPD aに格納されている。オペレーティングシステムOS a用の6つのファイルは、共通して使用されるファイルであるため、共通ボリュームPD cに格納される。

【0045】

一方、クライアント100 bについては、仮想ボリュームVD bに格納されているファイ

50

ルのうち、アプリケーション1、アプリケーション3、文書ファイルbは、実際には固有ボリュームPD bに格納されている。オペレーティングシステムOS b用の6つのファイルのうち、環境設定に依存しないファイルは共通ボリュームPD cに格納されている。環境設定を表すファイルCONF. SYS b, AUTO. BAT bは、クライアント100 bに固有のファイルとして扱われており、固有ボリュームPD bに格納される。

【0046】

固有ボリュームPD a、PD bおよび共通ボリュームPD cに格納すべきデータの分類については、種々の設定が可能である。例えば、デフォルトの環境設定を表すファイルCONF. SYS, AUTO. BATについても、クライアント100 aの固有のファイルとして扱い、固有ボリュームPD aに格納してもよい。別の設定として、クライアント100 bの環境設定を表すファイルCONF. SYS b, AUTO. BAT bも含め、リモートブートで使用するファイルを全て共通ボリュームPD cに格納するものとしてもよい。更に、別の設定として、クライアント100 a、100 bが共通して使用するアプリケーション1を共通ボリュームPD cに格納してもよい。

【0047】

C. 機能ブロック；

図4は計算機システムを構成する各装置内の機能ブロックを示す説明図である。本実施例では、図示する各機能ブロックは、それぞれコンピュータプログラムをインストールすることによりソフトウェア的に構成される。各機能ブロックは、ハードウェア的に構成することも可能である。

【0048】

計算機100の機能ブロックは、次の通りである。TCP/IPモジュール120は、ネットワークを介してIPパケットを通信する。SCSIモジュール122は、ストレージ装置300に対してデータの読み出し、書き込み用のコマンド送受信などを行う。iSCSIモジュール121は、SCSIコマンドとIPパケットの変換を行うことによって、SCSIコマンドをTCP/IPで送受信可能とする。ブートプログラム130は、計算機100でリモートブートを実現する。ブートプログラム130は、計算機100の初期化を行う機能、起動に必要なオペレーティングシステムなどのファイルをストレージ装置から読み込む機能、これらのファイルを起動してブートする機能などを実現することができる。ID入力モジュール131は、リモートブートを行う際に、ストレージ装置300に送信すべきユーザID、パスワードを入力する。本実施例では、これらのモジュールは、ROM103に記憶されているプログラムによって実現される。

【0049】

上述の各機能ブロックによってリモートブートが完了すると、計算機100には、ストレージ装置300から供給されたファイルによってオペレーティングシステム140が起動する。計算機100上の各処理に利用されるアプリケーション150は、ストレージ装置300から供給され、オペレーティングシステム140の上で稼働する。

【0050】

ノード装置200には、ネットワークを介してSCSIコマンドを授受するため、TCP/IPモジュール220、iSCSIモジュール221、SCSIモジュール222が備えられている。ノード装置200には、所定のオペレーティングシステムがインストールされており、このオペレーティングシステム上で稼働するアプリケーションによってキャッシュ制御モジュール230が実現されている。

【0051】

キャッシュ制御モジュール230は、物理ブロック管理テーブル234、キャッシュ管理テーブル232を参照して、ノード装置200におけるキャッシュを制御する。物理ブロック管理テーブル234とは、仮想ボリュームにおけるデータの格納場所、即ち仮想ブロックと、ストレージ装置300におけるデータの格納場所、即ち物理ブロックとを対応づけるテーブルである。キャッシュ管理テーブル232とは、物理ブロックと、ノード装置200におけるデータの格納場所、即ちキャッシュブロックとを対応づけるテーブルであ

10

20

30

40

50

る。これらのテーブルのデータ構造については後述する。物理ブロック管理テーブル 2 3 4 およびキャッシュ管理テーブル 2 3 2 を順次参照することにより、クライアントから指定された仮想ブロックに対応するデータが格納されているキャッシュブロックを特定することができる。

【0052】

ストレージ装置 3 0 0 には、ネットワークを介して S C S I コマンドを授受するため、T C P / I P モジュール 3 2 0、i S C S I モジュール 3 2 1、S C S I モジュール 3 2 2 が備えられている。ストレージ装置 3 0 0 には、所定のオペレーティングシステムがインストールされており、このオペレーティングシステム上で稼働するアプリケーションによってアクセス制御モジュール 3 3 0 が実現されている。

10

【0053】

アクセス制御モジュール 3 3 0 は、ユーザ管理テーブル 3 3 4、物理ブロック管理テーブル 3 3 2 を参照して、ディスク装置 3 1 0 のデータの読み書きを制御する。ユーザ管理テーブル 3 3 4 は、ユーザ I D と仮想ボリュームとを対応づけるテーブルである。物理ブロック管理テーブル 3 3 2 とは、仮想ブロックと物理ブロックとを対応づけるテーブルである。アクセス制御モジュール 3 3 0 は、計算機 1 0 0 の起動時に、計算機 1 0 0 からユーザ I D を取得する。このユーザ I D に基づいてユーザ管理テーブル 3 3 4 を参照することで、ユーザに対応する仮想ボリュームを計算機 1 0 0 に割り当てる。本明細書では、このように仮想ボリュームが特定された状態を「クライアント」と称する。

【0054】

アクセス制御モジュール 3 3 0 は、以後、クライアントから指定された仮想ブロックに対して、物理ブロック管理テーブル 3 3 2 を参照して、ディスク装置 3 1 0 のデータの読み書きを制御することができる。

20

【0055】

D. テーブルのデータ構造；

図 5 はユーザ管理テーブル 3 3 4 の構造を示す説明図である。ユーザ管理テーブル 3 3 4 は、ユーザ I D、パスワード、仮想ボリューム、属性、I P アドレスを対応づけて管理する。属性とは、仮想ボリュームに対する他のユーザのアクセスの可否を表している。「専有」は他のユーザのアクセスが禁止されていることを意味し、「共有」は許可されていることを意味する。I P アドレスは、ユーザがログインした計算機によって変動する。属性、I P アドレスは、ユーザ管理テーブル 3 3 4 から省略してもよい。

30

【0056】

図 6 は物理ブロック管理テーブル 3 3 2 の構造を示す説明図である。図の上方にデータ構造を示し、下方に仮想ブロックと物理ブロックとの対応関係を模式的に示した。先に説明した通り、物理ブロック管理テーブル 3 3 2 は、仮想ブロックと物理ブロックの対応関係を示すテーブルであり、仮想ボリューム単位で設けられている。図中には、仮想ボリューム V D a に対応するテーブルを例示した。このテーブルでは、仮想ボリューム V D a の仮想ブロック B L 0 ~ B L n に対して、実際にデータが格納されている物理ブロックが記憶されている。物理ブロックは、物理ボリュームとブロック番号との組み合わせで定義される。

40

【0057】

図に例示したテーブルは、下方に模式的に示した対応関係を表している。つまり、仮想ブロック B L 0 のデータは、共通ボリューム P D c のブロック B L 0 に格納されている。同様にして、仮想ブロック B L 1 ~ B L n の各データは、それぞれ共通ボリューム P D c または固有ボリューム P D a のいずれかのブロックに格納されている。このテーブルを参照することにより、クライアントが指定した仮想ブロックに対応するデータを、共通ボリューム P D c または固有ボリューム P D a から読み出すことが可能となる。物理ブロック管理テーブル 3 3 2 は、仮想ブロック V D b に対しても同様の形式でデータを記録している。仮想ブロック V D b に対しては、共通ボリューム P D c と固有ボリューム P D b のブロックが割り当てられることになる。共通ボリューム P D c のブロックに対しては、仮想ブ

50

ロック V D a、V D b の双方が対応づけられる可能性がある。かかる対応付けにより、共通ボリューム P D c のデータを複数のクライアントで共有することが可能となる。

【0058】

ノード装置 200 が備える物理ブロック管理テーブル 234 も同様の形式でデータを格納する。ノード装置は、物理ブロック管理テーブル 234 として、ストレージ装置 300 に格納されている全てのデータに関する情報を管理してもよいし、キャッシュされているデータに関する部分の情報のみを管理してもよい。

【0059】

図 7 はキャッシュ管理テーブル 232 の構造を示す説明図である。このテーブルは、図の右側に示す通り、ストレージ装置 300 における物理ブロックと、ノード装置 200 におけるキャッシュブロックとを対応づけている。物理ブロックは、物理ボリュームとブロック番号で特定される。

10

【0060】

図中の左側には、2つの仮想ボリューム V D a、V D b にそれぞれ対応する物理ブロック管理テーブル 234 a、234 b の内容を併せて例示した。ノード装置 200 は、物理ブロック管理テーブル 234 a、234 b およびキャッシュ管理テーブル 232 を参照することにより、仮想ブロックに対応するデータの格納場所を特定することができる。例えば、仮想ボリューム V D a の仮想ブロック B L 0 は、物理ブロック管理テーブル 234 a により、物理ボリューム P D c のブロック B L 0 に対応するデータであることが特定される。次に、キャッシュ管理テーブル 232 を参照すると、物理ボリューム P D c のブロック B L 0 は、キャッシュブロック B L 0 に格納されていることがわかる。

20

【0061】

図では、キャッシュブロックは、物理ボリュームと無関係に割り当てられる場合を例示した。ノード装置 200 においても、共通ボリューム P C c と固有ボリューム P C a、P C b が定義されているから、これらに振り分けてデータを格納してもよい。例えば、ストレージ装置 300 において共通ボリューム P D c に格納されているデータを、共通ボリューム P C c に格納し、固有ボリューム P D a、P D b に格納されているデータを、それぞれ固有ボリューム P C a、P C b に格納してもよい。

【0062】

別の態様として、固有ボリュームについてはクライアントを区別することなくデータを格納するようにしてもよい。例えば、ストレージ装置 300 において固有ボリューム P D a、P D b に格納されているデータを区別することなく、同一の固有ボリュームに格納するようにしてもよい。

30

【0063】

E. データ読出処理；

図 8 はデータ読み出し処理のフローチャートである。リモートブート時を例にとり、左側にクライアント 100 の処理、中央にノード装置 200 の処理、右側にストレージ装置 300 の処理を示した。

【0064】

ユーザが計算機の電源をオンにすると、I D 入力モジュール 131 によって、ユーザ I D、パスワード等の入力画面が提示される。ユーザが、ユーザ I D、パスワードを入力し、ログインすると、クライアント 100 は、I P アドレスとともにこれらの情報を、ノード装置 200 およびストレージ装置 300 に送信する（ステップ S 10）。ストレージ装置 300 は、この情報に基づいてユーザ管理テーブル 334 を参照し、ユーザの認証を行う（ステップ S 30）。正規なユーザであることが確認されると、ストレージ装置 300 は、ユーザ管理テーブル 334 に、クライアント 100 の I P アドレスを記録するとともに、ユーザに対応した仮想ボリューム名を通知する。この通知は、ノード装置 200 を介してクライアント 100 に送信される。以後、クライアント 100 からのデータの読み出し要求等は、この仮想ボリューム名に基づいて行われる。

40

【0065】

50

クライアント１００は、ＲＯＭに記録されているブートプログラム１３０を起動する（ステップＳ１１）。ブートプログラム１３０は、ブートに必要な初期化処理を行った上で、オペレーティングシステムのデータ読み出し要求をノード装置に対して送信する（ステップＳ１２）。

【００６６】

ノード装置２００は、読み出し要求を受信し、既にキャッシュ済みであるか否かを確認する（ステップＳ２０）。この確認は、図７で説明した通り、物理ブロック管理テーブル２３４とキャッシュ管理テーブル２３２を参照して行う。キャッシュ済みの場合は（ステップＳ２０）、キャッシュブロックから該当するデータを読み出し、クライアント１００に返信する（ステップＳ２５）。クライアント１００は、このデータを受信して起動処理を継続する（ステップＳ１３）。 10

【００６７】

要求されたデータがキャッシュされていない（以下、「未キャッシュ」と呼ぶ）場合（ステップＳ２０）、ノード装置は、ストレージ装置に対して、クライアント１００から要求された仮想ブロックがどの物理ブロックに格納されているかを問い合わせる（ステップＳ２１）。ストレージ装置３００は、この問い合わせを受けると、物理ブロック管理テーブル３３２を参照して、仮想ブロックが格納されている物理ブロックの物理ボリュームおよびブロック番号を、ノード装置２００に返信する（ステップＳ３１）。

【００６８】

ノード装置２００は、物理ブロック管理テーブル２３４に、ストレージ装置３００から受信した情報を格納し、クライアント１００から要求されたデータがキャッシュ済みか否かを、再度確認する（ステップＳ２２）。この時点では、物理ブロック管理テーブル２３４が更新されただけであり、ストレージ装置３００から新たにデータを受信した訳ではないが、キャッシュ済みと判断される場合が起こり得る。例えば、クライアント１００が、共通ボリュームに格納されているデータの読み出しを要求している場合である。共通ボリュームに格納されているデータは、既に他のクライアントの読み出し要求に応じて、ノード装置２００にキャッシュされている場合がある。新たに接続されたクライアント１００から要求された仮想ブロックについて、ストレージ装置３００から物理ブロックに関する情報を取得して初めて、キャッシュ済みのデータとの対応関係が判明した場合には、この時点で、キャッシュ済みと判断される。キャッシュ済みと判断した場合には（ステップＳ２ 30
２）、ノード装置２００は、キャッシュされているデータをクライアントに返信し（ステップＳ２５）、クライアントはこれを受信する（ステップＳ１３）。

【００６９】

物理ブロック管理テーブル２３４を更新したにもかかわらず、未キャッシュと判断した場合（ステップＳ２２）、ノード装置２００はストレージ装置３００にデータの読み出しを要求する（ステップＳ２３）。ストレージ装置３００は、この要求に応じて、共通ボリュームまたは固有ボリュームからデータを読み出し、ノード装置２００に返信する（ステップＳ３２）。

【００７０】

ノード装置２００は、ストレージ装置３００からデータを受信すると、これをキャッシュし、キャッシュ管理テーブル２３２を更新する（ステップＳ２４）。また、受信したデータをクライアントに返信し（ステップＳ２５）、クライアントはこれを受信する（ステップＳ１３）。 40

【００７１】

上述の処理において、ノード装置２００において、ステップＳ２０で未キャッシュと判断されたデータが、物理ブロック管理テーブル２３４を更新することにより、ステップＳ２２でキャッシュ済みと判断される具体例を示す。図７において、先に仮想ボリュームＶＤｂ、ＢＬ０のデータがキャッシュブロックＢＬ０に既にキャッシュ済みであり、クライアントから仮想ボリュームＶＤａ、ＢＬ０のデータについて読み出し要求を受けた場合を考える。ノード装置２００は、この時点では、物理ブロック管理テーブル２３４ａの仮想ブ 50

ロック B L O に対応するデータを管理していないため、要求されたデータは未キャッシュと判断する。その後、ストレージ装置 3 0 0 から、物理ブロック管理テーブル 2 3 4 a の 1 行目のデータ、即ち、仮想ブロック B L O が物理ディスク P D c、ブロック B L O に格納されている情報を取得する。このデータを取得すると、ノード装置 2 0 0 は、更新後の物理ブロック管理テーブル 2 3 4 a、およびキャッシュ管理テーブル 2 3 2 を参照することにより、要求されたデータがキャッシュブロック B O O にキャッシュ済みであると判断することができる。

【 0 0 7 2 】

以上の処理を繰り返し行うことにより、クライアント 1 0 0 は、ブートに必要なファイルを取得し、リモートブートを完了することができる。図 8 ではリモートブート時の読み出し処理を示したが、リモートブートと無関係に通常のデータ読み出し時にも、データ読み出し要求（ステップ S 1 2）からデータ受信（ステップ S 1 3）の処理が同様にして行われる。

【 0 0 7 3 】

以上で説明した実施例の計算機システムでは、リモートブート用のデータは、複数のクライアントに共有される共通ボリュームに格納され、ノード装置においても複数のクライアントに共通のデータとしてキャッシュされる。従って、これらのデータについて複数のクライアントからのアクセスが集中しても、データ読み出し時におけるディスク装置のヘッド移動量を抑制することができ、読み出し速度低下を抑制することができる。

【 0 0 7 4 】

F. 変形例；

（１） 実施例では、ノード装置 2 0 0 は、クライアントから要求されたデータが未キャッシュと判断された場合に（図 8 のステップ S 2 0）、ストレージ装置 3 0 0 に物理ブロックに関する情報の問い合わせを行う場合を例示した。これに対し、ストレージ装置 3 0 0 で管理している物理ブロック管理テーブル 3 3 2 の全データを、予めノード装置 2 0 0 に送信しておいてもよい。こうすることにより、ノード装置 2 0 0 は、ストレージ装置 3 0 0 への問い合わせ（図 8 のステップ S 2 1）を行うまでなく、要求されたデータがキャッシュ済みかを判断することが可能となる。

【 0 0 7 5 】

（２） 実施例では、ストレージ装置 3 0 0 は、物理ブロックに関する情報の問い合わせを受けたノード装置 2 0 0 に対してのみ、物理ブロック管理テーブルの情報を返信する場合を例示した（図 8 のステップ S 3 1）。これに対して、ストレージ装置 3 0 0 は、全てのノード装置に対して、物理ブロック管理テーブルの情報をマルチキャストしてもよい。こうすることにより、ユーザが別のノード装置に接続されている計算機 1 0 0 で再ログインした場合でも、ノード装置は、ストレージ装置 3 0 0 から得られた情報を有効活用して、キャッシュ済みか否かの判断を迅速に行うことが可能となる。

【 0 0 7 6 】

（３） 実施例では、ノード装置が設けられている計算機システムを例示した。これに対し、本発明は、ストレージ装置に計算機が直接、接続されている計算機システムに適用することも可能である。かかる計算機システムでは、図 8 におけるノード装置の処理が省略される。ストレージ装置は、クライアントからの要求（ステップ S 1 2）に応じて、物理ブロック管理テーブル 3 3 2 を参照し、データの返信を行えばよい（ステップ S 3 2、S 1 3）。かかる処理でも、複数のクライアントに共通するデータを共通ボリュームで管理することにより、読み出し速度の低下を抑制するという効果を得ることができる。

【 0 0 7 7 】

以上、本発明の種々の実施例について説明したが、本発明はこれらの実施例に限定されず、その趣旨を逸脱しない範囲で種々の構成を採ることができることはいうまでもない。

【 0 0 7 8 】

【 発明の効果 】

本発明によれば、共通ボリュームと固有ボリュームに区分けして、データを格納すること

10

20

30

40

50

により、複数のクライアントから共通して読み出されるデータを共通ボリュームでまとめて管理することができる。共通ボリュームにまとめて格納されたデータについては、複数のクライアントからのアクセスが集中しても、ヘッドの移動量を抑制することができ、データの読み出し速度低下を抑制することができる。

【図面の簡単な説明】

【図1】実施例としての計算機システムの構成を示す説明図である。

【図2】データの格納および読出状態を模式的に示す説明図である。

【図3】データの格納例を示す説明図である。

【図4】計算機システムを構成する各装置内の機能ブロックを示す説明図である。

【図5】ユーザ管理テーブル334の構造を示す説明図である。

10

【図6】物理ブロック管理テーブル332の構造を示す説明図である。

【図7】キャッシュ管理テーブル232の構造を示す説明図である。

【図8】データ読み出し処理のフローチャートである。

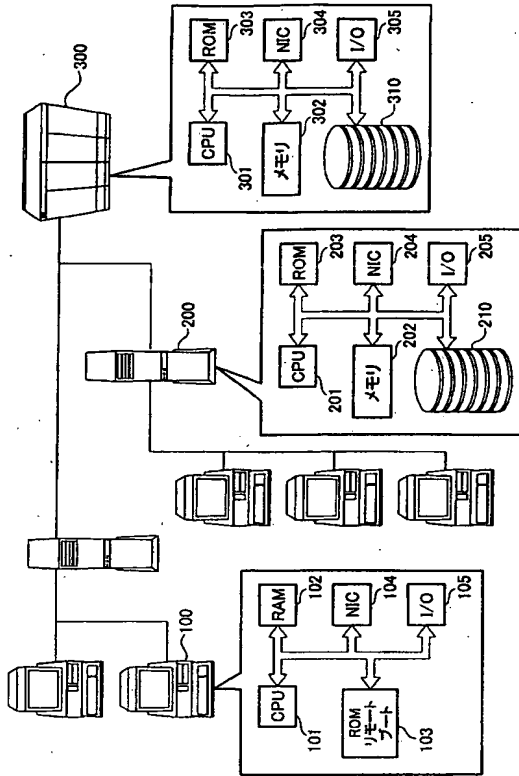
【符号の説明】

- 100、100a、100b…クライアント
- 101、201、301…CPU
- 102、202、302…メモリ(RAM)
- 103、203、303…ROM
- 104、204、304…NIC
- 105、205、305…I/O
- 120、220、320…TCP/IPモジュール
- 121、221、321…iSCSIモジュール
- 122、222、322…SCSIモジュール
- 130…ブートプログラム
- 131…ID入力モジュール
- 140…オペレーティングシステム
- 150…アプリケーション
- 200…ノード装置
- 210、310…ディスク装置
- 230…キャッシュ制御モジュール
- 232…キャッシュ管理テーブル
- 234、234a、234b…物理ブロック管理テーブル
- 300…ストレージ装置
- 330…アクセス制御モジュール
- 332…物理ブロック管理テーブル
- 334…ユーザ管理テーブル

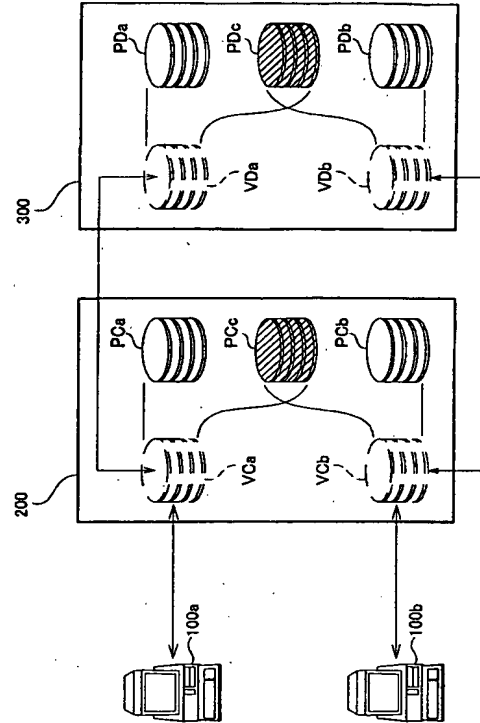
20

30

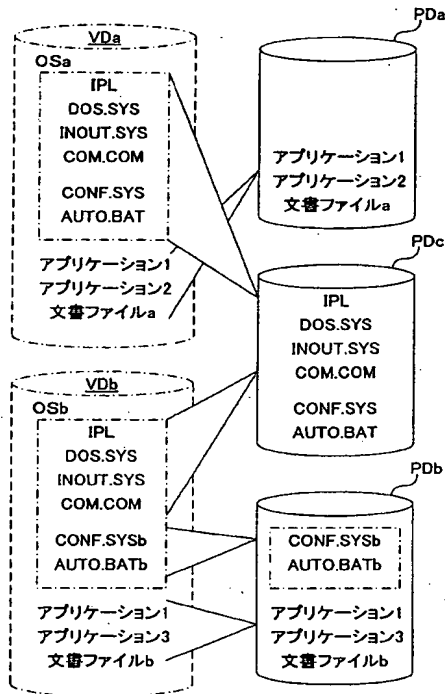
【図 1】



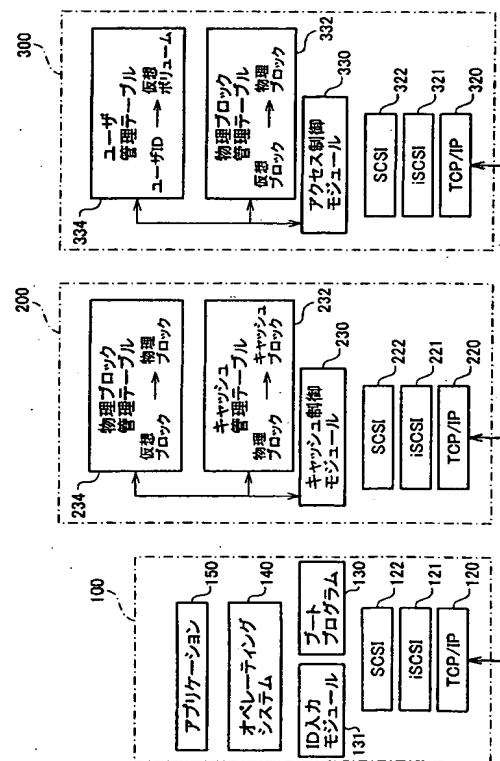
【図 2】



【図 3】



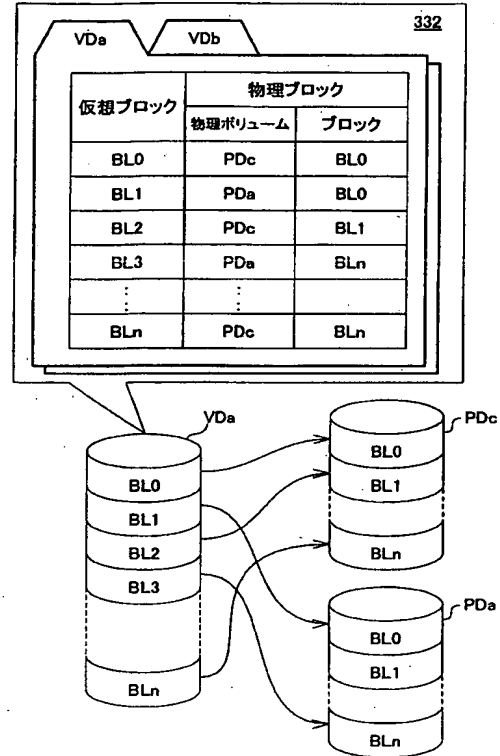
【図 4】



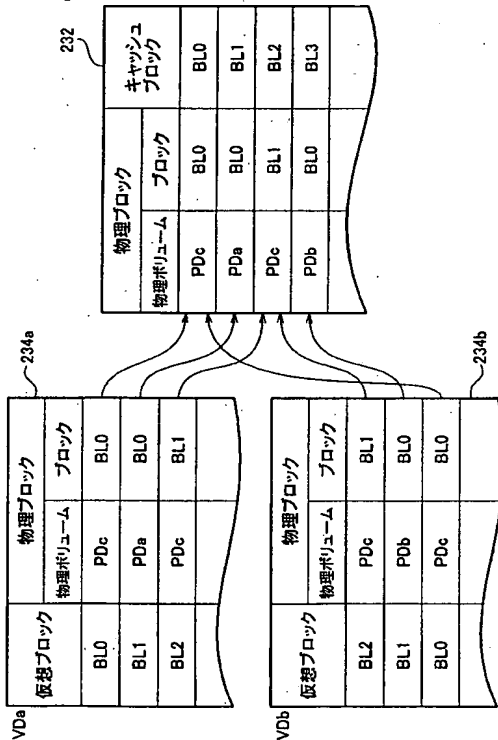
【図 5】

ユーザID	パスワード	仮想ボリューム	属性	IPアドレス
ユーザA	*****	VDa	専有	IPa
ユーザB	*****	VDb	専有	
ユーザC	*****	VDe	共有	

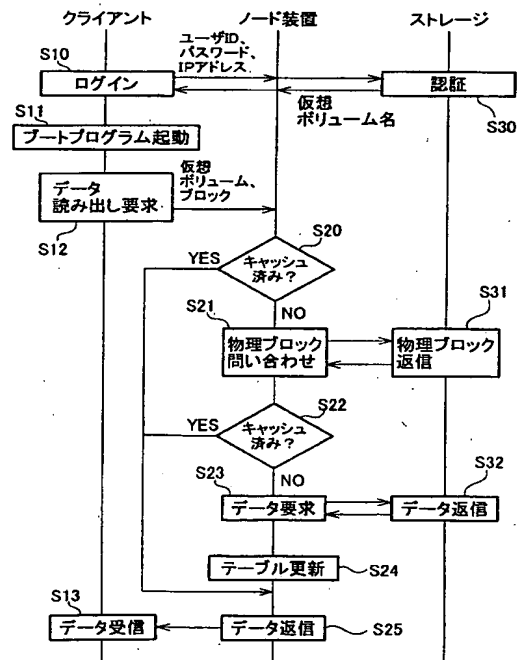
【図 6】



【図 7】



【図 8】



フロントページの続き

(72)発明者 橋本 尚

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア事業部内

Fターム(参考) 5B065 BA01 CA11 CA14 ZA01

5B082 FA05 HA08